DOCUMENT RESUME

ED 194 607 TM 800 738

TITLE Massachusetts Assessment of Basic Skills 1979-80. Summary Report. Listening, Speaking.
INSTITUTION Massachusetts State Dept. of Education, Boston. Bureau of Research and Assessment.
PUB DATE Sep 80
NOTE 66p.

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS *Educational Assessment; Grade 12; High Schools; *Listening Skills; *Minimum Competencies; *Speech Skills; *State Programs; Student Evaluation
IDENTIFIERS *Massachusetts

ABSTRACT

Statewide information about the performance of Massachusetts' grade 12 students in listening and speaking is presented for the 1979-80 school year. A total of 2,207 students were tested by listening to material on a tape recording and then answering multiple-choice questions about what they heard. The performance of students in listening is presented for each item and is summarized across all items. Performance is also summarized by type of listening skill and by type of listening situation. A total of 691 students were tested in two parts by: (1) two teachers independently rating a student's speaking performance from their observations during normal classroom activity; and (2) a trained rater assessing the student's speaking performance in a one-on-one situation during several specific speaking tasks. The performance of students in speaking is presented for each teacher observation and each one-on-one speaking rating, as well as a summary of all ratings. In addition, results from background questionnaires which included questions about general school and student characteristics are presented in an effort to amplify and explain the results of the skills assessments. (RL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED194607

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Massachusetts Assessment of Basic Skills 1979-80

TM800738

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
C. DiBella
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Summary Report

Massachusetts Department of Education

MASSACHUSETTS BOARD OF EDUCATION

Ann H. McHugh, Chairperson
James L. Green, Vice Chairperson

John Anthony
John W. Bond
Millie Clements
Howard A. Greis
Charles T. Grisby
Mary Ann Hardenbergh
Armando Martinez
Joseph C. Mello, Jr.
Edwin M. Rossman
Donald R. Walker

Gregory R. Anrig
Commissioner of Education, Secretary

Ex Officio Member:

Laura Clausen, Chancellor
Board of Higher Education

BUREAU OF RESEARCH AND ASSESSMENT

Allan S. Hartman, Director
Tracy Libros
Leslie S. May
Matthew H. Towle
Dorothy Earle
Betty Hancock
Cathie Ridge

The Massachusetts Department of Education insures equal employment/ educational opportunities/affirmative action, regardless of race, color, creed, national origin, or sex, in compliance with Title IX, or handicap in compliance with Section 504.

MASSACHUSETTS ASSESSMENT OF BASIC SKILLS 1979-80

SUMMARY REPORT

LISTENING • SPEAKING

Massachusetts Department of Education
Bureau of Research and Assessment
31 St. James Avenue
Boston, Massachusetts 02116

September 1980

Publication #12038 approved by Alfred C. Holland, State Purchasing Agent

FOREWORD

The 1979-80 Massachusetts Assessment of Basic Skills represents a continuation of the series of annual evaluations of the knowledge, understanding, skills, and attitudes of Massachusetts public school students. This year's program was specifically designed to be responsive to the State Board of Education's Policy on Basic Skills Improvement. The assessment this year was in listening and speaking. In combination with last year's assessment results for reading, writing, and mathematics, we now have a picture of how well high school seniors perform in all five basic skills areas.

About 2,900 twelfth-grade students from 49 high schools participated in the assessment program. This report describes the findings of the assessment as well as some interpretations of these findings by the subject area committees involved in advising the Department during the past year. To them, and to the many others whose involvement may not have been explicitly acknowledged, the Board of Education expresses its sincere appreciation.

The State Listening Test developed this year formed the basis for creating equivalent test forms for listening. As noted in Section 40.05 of the Regulations for Implementation of the Policy on Basic Skills Improvement, school districts choosing to use state tests will be provided with equivalent test forms by the Department of Education following approval of secondary-level basic skills plans. School districts who may be considering this option, therefore, are advised to obtain a copy of the development report which accompanies this summary in order to examine the technical properties of the test in more detail.

The Board has voted to delay the required implementation of the speaking component until no later than the 1984-85 school year. The development of equivalent State Speaking Tests for local district use is still in process. One primary goal in this process is to make these tests administratively convenient to use.

Meanwhile, until speaking assessment instruments are available, districts may still begin work in this important area. The Board of Education strongly encourages districts to review existing curriculum for speaking, establish speaking objectives, and begin focusing instruction toward those objectives.

It is our hope that this report will be widely read by both school and lay audiences. Even more than this, however, the Board of Education would like to see the information in this report used to bring about improvements in learning opportunities for all Massachusetts students.

Anne H. McHugh
Chairperson
Massachusetts Board of Education

TABLE OF CONTENTS

I. INTRODUCTION 1

Background 1
Instruments. 1
Test Administration. 2
Sampling 2
Reporting Results. 4

II. LISTENING. 7

Objectives and Instrument Development. 7
Listening Results. 8
Other Available Information. 22

III. SPEAKING 23

Objectives and Instrument Development. 23
Results of Teacher Observation and One-on-One Ratings 26
Other Available Information. 41

IV. BACKGROUND INFORMATION 43

Introduction 43
School Characteristics 43
Student Characteristics. 45
Other Available Information. 46

APPENDIX 47

ACKNOWLEDGMENTS. 55

LIST OF EXHIBITS AND TABLES

TABLE 1. Percentage of Students in the Listening and Speaking Samples by Categories 5

EXHIBIT 1. Listening Objectives 8

TABLE 2. Distribution of Listening Scores 10

EXHIBIT 2. Average Percentage of Items Answered Correctly for the Total Listening Test and by Each Listening Skill 11

EXHIBIT 3. Average Percentage of Items Answered Correctly for the Total Listening Test and by Each Listening Situation 13

EXHIBIT 4. Percentage of Students Choosing Each Response for Each Listening Item. 14

TABLE 3. Average Percentage of Items Answered Correctly on the Total Listening Test for All Students and by Various Groups of Students 21

EXHIBIT 5. Speaking Objectives. 24

EXHIBIT 6. Average Rating for the Teacher Observation Measure Overall and by Each Speaking Dimension 28

TABLE 4. Percentage of Students Receiving Teacher Observation Ratings at Each Level for Each Dimension 29

TABLE 5. Distribution of Total Teacher Observation Scores 30

EXHIBIT 7. Average Rating for the One-on-One Measure Overall and by Each Speaking Dimension. 32

EXHIBIT 8. Average Rating for the One-on-One Measure Overall and by Each Speaking Task 33

TABLE 6. Percentage of Students Receiving One-on-One Speaking Ratings at Each Level for Each Dimension of the Description Task 34

TABLE 7. Percentage of Students Receiving One-on-One Speaking Ratings at Each Level for Each Dimension of the Emergency Task 35

TABLE 8. Percentage of Students Receiving One-on-One Speaking Ratings at Each Level for Each Dimension of the Sequence Task. 36

TABLE 9. Percentage of Students Receiving One-on-One Speaking Ratings at Each Level for Each Dimension of the Persuasion Task. 37

TABLE 10. Distribution of Total One-on-One Scores 39

TABLE 11. Average Teacher Observation Ratings and Average One-on-One Ratings for All Students and by Various Reporting Groups 40

I. INTRODUCTION

Background

For the sixth consecutive year the Massachusetts Department of Education conducted a statewide educational assessment. The 1979-80 assessment focused on the basic skills of listening and speaking. The assessment fulfills two purposes. First, it provides statewide information about the performance of 12th graders in listening and speaking. Second, it provides information necessary for developing equivalent assessment instruments in listening and speaking that may be used by school districts in future years. Both purposes address requirements of the Massachusetts Basic Skills Improvement Policy.

The scope of the assessment was established by the listening and speaking objectives identified in the Basic Skills Improvement Policy. These objectives focus on skills essential for everyday functioning in school, employment and community settings. They are minimal competencies that all students should master before leaving high school. Thus, the goal of the assessment is to measure the most important and basic listening and speaking skills, not the full range of skills within these areas.

The assessment was designed and implemented by the Education Commission of the States (ECS), a private, nonprofit organization located in Denver, Colorado, under contract to the Massachusetts Department of Education. ECS worked closely with the Department of Education on all phases of the assessment. Critical aspects of the project were reviewed by various state groups and committees including the State Board of Education, the Bureau of Equal Educational Opportunity, the Basic Skills Improvement Advisory Committee, the Communication Skills Advisory Committee and the Listening/Speaking Skills Advisory Committee. The purpose of these reviews was to assure that the assessment reflected the intent of the Basic Skills Improvement Policy and that instruments were reliable, valid, free of bias and could be easily utilized by school districts in the future.

Instruments

For the listening assessment, groups of students listened to material on a tape recording and then answered multiple-choice

questions about what they heard. The instrument is unique in that the passages students listened to represent typical spoken language such as news stories, public service announcements, commercials and conversations. Most existing listening tests use typical written material, read aloud.

The speaking assessment had two parts. For the first part, two teachers who had a student in class independently assessed the student using a set of speaking rating scales. The teachers based their ratings of a student on their observations of the student's typical speaking performance during normal classroom activity over a period of time. For the second part, a trained rater assessed the student's speaking performance in a one-on-one situation. The rater gave the student several specific speaking tasks and then rated the student's performance on each task using the same speaking rating scales.

In addition to the assessment of listening and speaking performance, background questionnaires were administered to students, teachers and principals. They included questions about general school and student characteristics and questions about specific factors that might influence listening and speaking performance. This information was collected to provide baseline information prior to the implementation of the Basic Skills Improvement Policy and to help explain the performance of students on the listening and speaking assessment.

Test Administration

The listening and speaking assessment was conducted during March and April 1980. Forty-nine secondary schools were randomly selected to participate in the assessment. In each school, 12th graders were randomly selected to be assessed. No student was assessed in both areas, and testing was completed during a single school day.

The testing was conducted by a test administrator hired by ECS. The test administrator was assisted by a school coordinator identified by each participating school. The test administrator was primarily responsible for conducting the one-on-one speaking assessment. The school coordinator was primarily responsible for conducting the listening assessment.

Sampling

The assessment used a sophisticated sampling procedure to select students for participation in the assessment. This sample was designed to be representative of high school seniors from four

kind-of-community (KOC) categories. These KOC categories are defined as follows:

- BIG CITIES: communities designated as central cities according to the 1970 U.S. Census. Examples of communities classified as big cities (although not necessarily in the sample) include Boston, Brockton, Haverhill, Holyoke, Lawrence and Worcester.

- INDUSTRIAL SUBURBS: suburbs of central cities with below-average family incomes (below $16,000) and with more than 20% commercial and 7% industrial land use. Examples include Cambridge, Chelsea, Medford, Quincy, Peabody, Waltham, Lee and Monson.

- RESIDENTIAL SUBURBS: suburbs of central cities other than Boston with above-average family incomes (above $16,000) and all non-industrial suburbs of Boston. Examples include Arlington, Braintree, Framingham, Scituate, Walpole, Andover, Dover, Milton, Newton, Chelmsford, Shrewsbury and Wilbraham.

- OTHER COMMUNITIES: resort towns in the Cape Cod and Berkshire areas, non-industrial suburbs of cities other than Boston with below-average family incomes, towns with below-average family income, and communities with a population of less than 2,500. Examples include Barnstable, Chatham, Yarmouth, Stockbridge, Bridgewater, Hadley, Oxford, Stoughton, Newburyport, Taunton, Ayer, Kingston, Orange, Webster, Ashfield, Dunstable and Plympton.

Within each KOC category, the sample was selected in two stages. First, schools were randomly selected. The chance of a school being selected depended on the number of seniors in the school. Procedures were followed to ensure adequate representation of minorities, areas of the state and sizes of schools. Secondly, students within each of the selected schools were randomly drawn and then randomly assigned to either the listening assessment or the speaking assessment.

Students were exempted from testing who are non-English speaking or handicapped (physically, mentally or emotionally) in such a way that they could not respond to the assessment. This is in accordance with Massachusetts Department of Education guidelines for its statewide assessments.

The sampling plan called for the random selection of 12 schools from within each KOC. Within each school 60 twelfth graders were randomly selected to participate in the listening assessment and 18 to participate in the speaking assessment. The plan assumed that the sample size would be reduced by approximately one-third due to ineligibility and absenteeism. In the actual assessment one additional school was added to make a total of

49 participating schools.[1] A total of 2,207 students were tested in the listening assessment and 691 students in the speaking assessment. Table 1 describes the composition of the samples for the listening and speaking assessments in terms of KOC, sex and racial/ethnic background. In the analysis of results the responses of students within each of the sampling groups were weighted according to their actual proportions in the state population. For example, responses of students from residential suburbs were weighted greater than responses of students from big cities because proportionally there are more of these students statewide.

Reporting Results

The listening results are reported in Chapter II. The performance of students in listening is presented for each item and is summarized across all items. Performance is also summarized by type of listening skill and by type of listening situation.

The speaking results are reported in Chapter III. The performance of students in speaking is presented for each teacher observation and each one-on-one speaking rating. Performance is also summarized across all teacher observation ratings, across all one-on-one ratings and in additional ways.

Results of the listening and speaking assessments include interpretations of and comments about the data offered by the Communication Skills and Listening/Speaking Skills Advisory Committees.

The results from student, teacher and principal background questionnaires are summarized in Chapter IV. These results help to amplify and explain the results of the skills assessments.

No individual student or school is identified in the results. Thus, the results reflect the average level of performance of students across the state. They do not indicate the level of performance of individual students or schools.

---

[1] One school was added to complement a school that was selected but did not have the number of seniors required by the sampling plan.

More in-depth information about the methodology and the results of the 1979-80 Basic Skills Assessment is available in the technical report. A copy of this report has been sent to each Superintendent of Schools in the state. Limited numbers of copies are also available from the Bureau of Research and Assessment and the Regional Educational Centers of the Department of Education.

TABLE 1. Percentage of Students in the Listening and Speaking Samples by Categories

	Listening	Speaking
KIND OF COMMUNITY		
Big Cities	25%	25%
Industrial Suburbs	25	25
Residential Suburbs	25	26
Other	25	25†
SEX		
Male	48	49
Female	51	50
Unknown	1	1
RACIAL/ETHNIC BACKGROUND		
Black	5	5
Hispanic	3	3
White	84	87
Other	2	1
Unknown	6	4

†Because of rounding error, percentages might not total 100%.

II. LISTENING

Objectives and Instrument Development

The listening assessment was designed to measure the 11 listening objectives identified by the Massachusetts Basic Skills Improvement Policy. (See Exhibit 1.) Some of the objectives deal with general listening skills that apply to all listening situations -- for example, Objective A-1, Recognize Words and Phrases. Other objectives deal with specific listening situations -- for example, Objective C-1, Understand and Respond to Survival Words Used in Emergency Situations.

To assess the listening objectives, a test was developed that includes a variety of listening passages and asks a variety of listening questions. The passages reflect the objectives that deal with specific listening situations. The questions reflect the objectives that deal with general listening skills. The test measures the objectives globally; it does not measure each individual objective. Thus, the instrument is a general descriptive test of listening ability, not a diagnostic test.

The listening test consists of six passages and 22 multiple-choice items (3 to 4 per passage). The passages include:

- A news story
- A public service announcement
- a conversation
- A commercial
- A telephone conversation
- A teacher's announcement

All passages and items were field tested prior to final selection and were judged by the review committees to be culturally fair and balanced with respect to sex, racial/ethnic and economic representation.

The listening test was designed to be virtually self-contained on a cassette tape. Students were provided with booklets that included instructions and the multiple-choice questions. Students heard the instructions, then a passage and then questions about the passage. Questions and response options were read aloud to minimize the effect of a respondent's reading ability on listening performance. After each item was read, there was a pause; this allowed students plenty of time to answer the question. They did so by marking their responses on an answer sheet.

EXHIBIT 1. Listening Objectives

A. BASIC LISTENING SKILLS

1. Recognize words and phrases used by the speaker
2. Indicate why the speaker can or cannot be understood

B. UNDERSTANDING WHAT YOU HEAR

1. Understand spoken words and ideas
2. Identify and understand main ideas
3. Associate important details with main ideas
4. Understand descriptions of events and experiences
5. Understand speaker's purpose

C. USING WHAT YOU HEAR

1. Understand and respond to survival words used in emergency situations
2. Summarize information and draw conclusions
3. Recognize when words and phrases are used to convince or persuade
4. Follow straightforward directions

Listening Results

Interpreting Results

The basic measure of listening performance is the percentage of students responding correctly to a given item. In addition, percentages are provided for each incorrect choice. Average, or mean, percentages are reported for groups of items. These groups include: the overall score (all items in the test), items grouped by listening skill and items grouped by listening situation. Students who responded to less than half of the items on the test are excluded from the results.

The listening results are based on the performance of a sample of 12th graders. They are estimates of the "true" performance that would have been obtained if all 12th graders in Massachusetts had responded to the items. Statistical computations, standard errors, indicate that these estimates are most likely within 2 percentage points of the "true" performance of the population. For example, if the results indicate that the average score of students from the sample is 85%, one can feel confident that the "true" score of students is between 83% and 87%. Thus, the reader should not overemphasize the importance of small differences in scores.

The passages and items that make up the listening test are only a sample of all possible listening passages and items that might have been included. Thus, the reader should be cautious in making generalizations about listening results.

Overall Listening Results

On the average, students correctly answered 88% of the items on the listening test.[2] The average score is high, which is to be expected in a basic skills test.

The average percentage correct for the total test, 88%, is equivalent to about 19 out of 22 items correct. Table 2 shows the distribution of students' scores on the listening test. It indicates the percentage and the cumulative percentage of students obtaining each possible raw score, from 0 to 22. The distribution shows that the largest percentage of students got scores of 19 or higher, and that only 10% of the students got scores of 16 or lower.

In addition to the overall results, it is possible to look at the listening results in several other ways. The following sections examine the listening results in terms of listening skills, listening situations, individual items and various groups of students.

---

[2]The standard error for the listening test is .5%.

TABLE 2. Distribution of Listening Scores

Raw Score	Percentage of Students	Cumulative Percentage
0	0	0
1	--	--
2	0	--
3	0	--
4	0	--
5	0	--
6	--	--
7	--	--
8	--	--
9	--	1
10	--	1
11	--	1
12	1	2
13	1	3
14	2	4
15	3	7
16	3	10
17	6	16
18	9	25
19	15	40
20	21	61
21	24	85
22	15	100

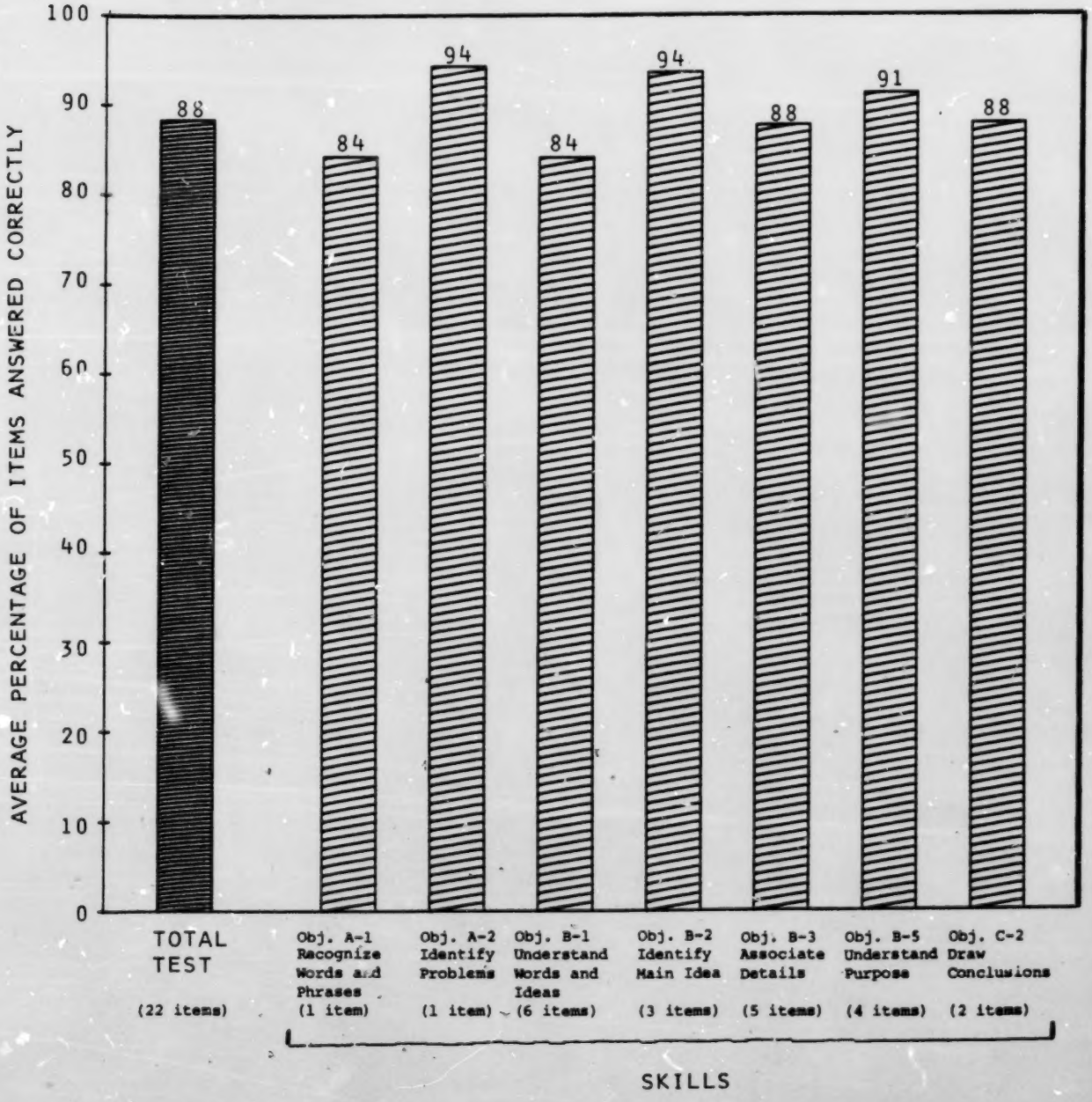
--Dashes are used for percentages that are less than .5.

Results by Listening Skills

Exhibit 2 shows the average percentages correct for the total test and for items grouped by the general listening skills in the objectives.

As can be seen from the graph, the percentages are high for all skills. Items related to two skills, Recognize Words and Phrases, and Understand Words and Ideas, appear relatively harder for students than the other items. Average percentages correct for three skills, Identify Problems, Identify Main Idea, and Understand Purpose, are above 90%, and these items appear to be very easy for students. Items related to the remaining two skills, Associate Details and Draw Conclusions, fall in between the two extremes, but students still perform well on these items.

EXHIBIT 2. Average Percentage of Items Answered Correctly for the Total Listening Test and by Each Listening Skill



NOTE: Total test performance is the average percentage of students answering correctly each of the 22 test items. It is equal to the weighted average of the scores for each of the six skills (i.e., each skill score is weighted by the number of items in the skill).

It should be noted, however, that there are very few items for each listening skill. This makes it inappropriate to place too much emphasis on the relative difficulty of the different skills.

Results by Listening Situations

Exhibit 3 shows average percentages correct for the total test and items grouped by the specific listening situations in the objectives. Again, it should be noted that there are few passages and items for each of the situations, making it difficult to generalize.

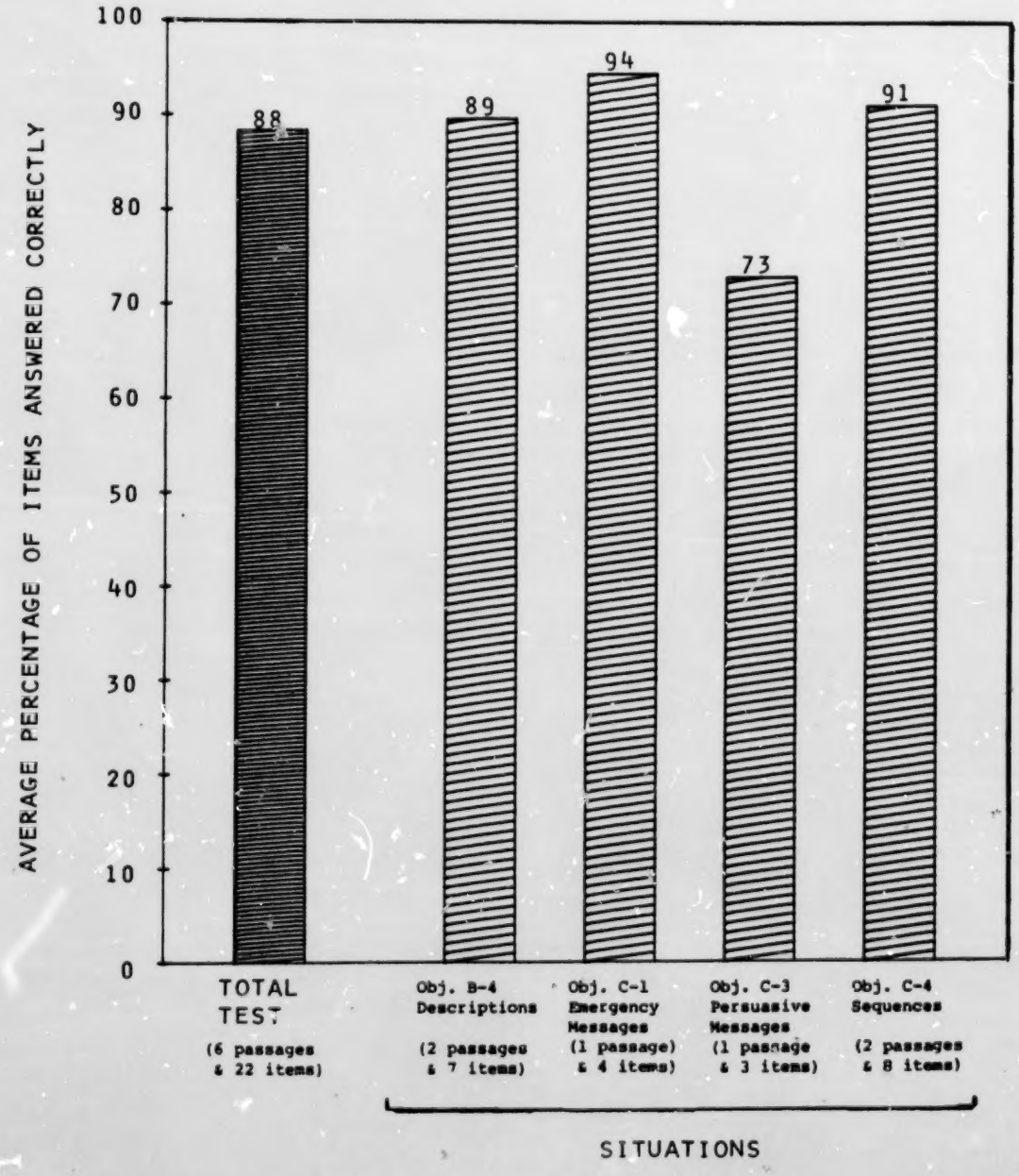
Students' performance is high for all situations except Persuasive Messages. For this situation, students heard a commercial and then responded to three items. This particular passage and items appear to be considerably more difficult for students than the other passages and their respective items. This may be because persuasive messages pose particularly difficult listening situations. It may also be accounted for by the difficulty level of one of the items (item 7) for the persuasive passage, which requires students to draw a very high-level inference about the material. The percentage of students who answered item 7 correctly is far lower than for all other items in the test. (See Exhibit 4.) In a previous Massahusetts Basic Skills Assessment of reading, similar results were obtained. Students' performance was noticeably lower on items associated with persuasive reading passages than on items for other types of reading material.

Item-by-Item Results

Item-by-item results are given in Exhibit 4. The objective for each passage is given below the passage, and the objective for each item is given to the left of each item number. Correct responses for each item are indicated by asterisks. The column to the right of the response choices indicates the percentage of students who chose that response. For example, the results for item 1 indicate that 90% of the students chose the correct response.

The item-by-item results give the most detailed information about performance of students on the listening test. They show the specific passage and item to which the students responded. By examining appropriate groups of passages and items, it is possible to get an idea of the tasks that are a part of the various summaries by listening skill, listening situation and the overall test.

EXHIBIT 3. Average Percentage of Items Answered Correctly for the Total Listening Test and by Each Listening Situation



NOTE: Total test performance is the average percentage of students answering correctly each of the 22 test items. It is equal to the weighted average of the scores for each of the four situations (i.e., each situation score is weighted by the number of items in the situation).

EXHIBIT 4. Percentage of Students Choosing Each Response for Each Listening Item

NEWS STORY

You are about to hear a recording of a news story. After the story you will be asked several questions about what you heard.

Do not take notes while you are listening to the recording.

Do not turn the page until told to do so. Now listen carefully.

The artistic process is a complex one -- take song writing for example. You picture the songwriter, sitting alone at the piano, composing some melodic work of art. But a Philadelphia firm has come up with a way of making that songwriter's job somewhat easier. Charles Trois of Tenel Industries has devised a computer system which he claims can help a songwriter pick a hit melody. To compute a song's hit potential, the melody is fed into the computer and within ten minutes the song is rated on a scale of one to ten. Trois explains his rating system as such: A song with a score of one to three will be a flop; a four means it's a possible hit; songs rated five and six could make the charts, and a seven might make the top ten. Melodies carrying ratings of eight and nine will be number one hits while a ten should become a classic. Trois claims that his computer has an eighty-two percent success rate in picking hits. However, there's only one problem with the machine -- it can only handle pop and soft rock songs. Hard rock just won't compute....

OBJ. B-4

			Percentage†
OBJ. B-2	1.	What is the main idea of this story?	
		A. Some songs are very hard to write.	2
		*B. Computers are now being used to pick hit songs.	90
		C. Computers can be programmed to write music.	6
		D. Most songs written will not be classics.	2
OBJ. B-1	2.	What can Trois' machine handle?	
		*A. Pop and soft rock	95
		B. Hard rock	2
		C. Classical music	2
		D. All time favorites	1
OBJ. A-1	3.	This story says that "a Philadelphia firm has come up with a way of making the songwriter's job somewhat easier." What does firm mean in the story?	
		A. Computer	6
		B. Inventor	9
		C. Lawyer	1
		*D. Company	84
OBJ. B-3	4.	How is a song rated?	
		A. By asking well known songwriters	1
		B. By singing the song for an audience	1
		C. By making the computer play the song	2
		*D. By feeding the melody into the computer	96

* Asterisks are used to indicate correct response.

† Because of rounding error and/or nonresponse, percentages might not total 100%.

--- Dashes are used for percentages that are less than .5.

EXHIBIT 4 *(Continued)*

COMMERCIAL

You are about to hear a recording of a commercial. After the commercial you will be asked several questions about what you heard.

Now listen carefully.

Enjoy delicious fried chicken right at home this week-end...and save money.. with a fresh whole fryer from Food King. Roast a whole fryer on the barbecue...split a fryer and broil the halves...cut up a fryer for fried chicken. Save on USDA Grade "A" whole fresh fryers, 10¢ off our regular low price per pound.

OBJ. C-3

			Percentage
OBJ. B-5	5.	What is the main purpose of this announcement?	
	A.	To convince people that Food King sells only top quality chicken	8
	B.	To give people information about food prices	2
	C.	To tell people how they can cook chicken	8
	*D.	To persuade people to buy chicken at Food King	83†
OBJ. B-1	6.	What did the announcement say about the price of chicken at Food King?	
	*A.	Chicken is 10¢ off per pound.	78
	B.	Chicken is 10% off.	10
	C.	Chicken is 20¢ off per pound.	3
	D.	Chicken is at its every day low price.	8†
OBJ. B-3	7.	Why should people buy chicken now at Food King?	
	A.	Because chicken is delicious when fried	3
	B.	Because Food King has fresh fryers	36
	*C.	Because chicken is on sale	56
	D.	Because Food King fryers are cut up	4†

EXHIBIT 4 *(Continued)*

TELEPHONE CONVERSATION

You are about to hear a recording of a telephone conversation. After the conversation you will be asked several questions about what you heard.

Now listen carefully.

Child: Hello?

Adult. Hello. I'd like to speak to Mrs. Kaplan.

Child: She's not here.

Adult: Oh, (pause) do you expect her back soon?

Child: No, she's out of town. She'll be back on Friday.

Adult: Ok, would you give her a message?

Child: Yeah, wait a minute. I've got to get something to write with. (Pause) Ok.

Adult: This is Mrs. Benton. I'm calling to remind your mother of the Neighborhood Action Group meeting this Saturday night. Tell her we'll need the box of pencils because we'll be voting. Also, she promised to give me a ride so I'll be expecting her to pick me up. Oh, yes, one more thing, we've got to be at the meeting at least 10 minutes early. Did you get all of that?

Child: Yeah, I got it. Do you want Mom to call you?

Adult: Yeah, I think that would be a good idea. I might think of something else before then. (This last sentence is spoken softly and mumbled.)

Child: What?

Adult: I said I might think of something else before then.

Child: Oh, Ok.

Adult: Thanks so much.

Child: Ok, bye.

Adult: Goodbye.

OBJ. C-4

			Percentage
OBJ. A-2	8.	Why did Mrs. Benton have to repeat part of her message?	
		A. Because there was a loud noise in the background	--
		B. Because she started to cough	1
		*C. Because she spoke too softly and mumbled	94
		D. Because she talked too fast	5
OBJ. C-2	9.	What will Mrs. Kaplan probably do as soon as she returns home?	
		A. Go to the meeting	1
		B. Go to her office for messages	8
		C. Type up the ballot for voting	1
		*D. Call Mrs. Benton	90
OBJ. B-3	10.	Why should Mrs. Kaplan bring pencils to the meeting?	
		A. To get addresses of new people	3
		B. To have for taking attendance	1
		C. To write down the message	2
		*D. To have for voting	96†
OBJ. B-5	11.	Why did Mrs. Benton call Mrs. Kaplan?	
		*A. To remind Mrs. Kaplan of the Neighborhood Action Group meeting	98
		B. To find out about Mrs. Kaplan's trip	1
		C. To give Mrs. Kaplan directions to her house	--
		D. To encourage Mrs. Kaplan to vote at the Neighborhood Action Group election	1

EXHIBIT 4 *(Continued)*

TEACHER'S ANNOUNCEMENT

You are about to hear a recording of a teacher's announcement. After the announcement you will be asked several questions about what you heard.

Now listen carefully.

Now, if I can have your attention for a few minutes, I'd like to explain a new assignment. Please settle down so everyone can hear.

After we finish our current unit, we'll have one free class period before break. What I'd like to do with this time is to have a party, but this won't be an ordinary party. It will be a special party because we're going to invite people from the past. And each one of you will be one of those historical people. During the party, you'll be asked many questions about your life and what you did that was important, so this will require a little bit of preparation. Tomorrow we'll spend some time in class drawing up a list of names of famous, historical people. This way we'll get a lot of different ideas. Next, we'll put all of the names in a box and each draw one name. Then, before the party, you'll have to learn about the person whose name you drew. (Pause) So, if there are no questions, I'll go on with today's topic.

OBJ. C-4

			Percentage
OBJ. B-5	12.	What is the main purpose of the teacher's announcement?	
	A.	To encourage students to read magazines and newspapers	2
	*B.	To explain a new assignment to the class	94
	C.	To remind students about school break	2
	D.	To urge students to always come to class well prepared	2
OBJ. B-1	13.	What will the class study before break?	
	*A.	People from the past	90
	B.	Political parties	3
	C.	Social customs	1
	D.	Current events	6
OBJ. B-2	14.	What are the students supposed to do?	
	*A.	Find out about a famous person	92
	B.	Go to a costume party	3
	C.	Write about a famous person	3
	D.	Collect pictures of famous people	1†
OBJ. B-1	15.	When will the students draw up a list of names?	
	A.	Before the end of the day	10
	*B.	Tomorrow	70
	C.	Next week	2
	D.	Before break	17†

EXHIBIT 4 *(Continued)*

PUBLIC SERVICE ANNOUNCEMENT

You are about to hear a recording of a Public Service Announcement. After the announcement you will be asked several questions about what you heard.

Now listen carefully.

Your City Parks and Recreation Department has something for you every season. Register now for the final fall session of Outdoor Nature Programs. A variety of wildlife, ecology, and conservation classes and activities are offered through November for the whole family. Young Explorers meet in the morning or after school and teens may become wildlife volunteers. Nature hikes and cross country ski trips are available for all ages and ability levels. For more information, call us at 472-4523.

OBJ. B-4

			Percentage
OBJ. B-5	16.	What is the main purpose of this announcement?	
	A.	To urge people to donate money to the Outdoor Nature Program	3
	B.	To inform people about current problems of ecology	3
	C.	To persuade people to protect birds and other wildlife	3
	*D.	To give people information about nature classes	90†
OBJ. B-1	17.	Who are the classes for?	
	A.	Residents only	5
	B.	Club members	3
	C.	Young Explorers only	7
	*D.	The whole family	85
OBJ. B-1	18.	Who is sponsoring the nature programs?	
	A.	The Young Explorers	4
	B.	A local conservation group	5
	C.	Wildlife volunteers	5
	*D.	The Parks and Recreation Department	86

EXHIBIT 4 *(Continued)*

CONVERSATION

You are about to hear a recording of a conversation. After the conversation you will be asked several questions about what you heard.

Now listen carefully.

Janice: I always enjoy this bus ride up to Vermont. It's so good to get away from the city.

Man on Bus: Yes, everything seems so much cleaner out here. I don't know, though, sometimes these buses make me feel a little queasy, all the shaking back and forth. They're always so stuffy. I'm beginning to feel a little dizzy right now. I think...(pause).

Janice: Hey, hey! Are you all right? Somebody, help me! He's passed out. Stop the bus!

Passenger: What's going on? What's the matter with him?

Janice: I think he's fainted. He said something about feeling dizzy. Could you help me? What should we do?

Passenger: Sure. Let's lower his head between his knees. Get his hands out of the way. Yeah, that looks good. He'll be all right in a minute.

Janice: Good, 'cause I don't know if I can hold him like this for long. When he said he felt dizzy, he really meant it!

Passenger: Yeah, he looks really pale. I think we'll be getting off the freeway at the next exit. Do you think you can hold him by yourself for a minute while I go ask the driver?

Janice: Yes, I can hold him a little bit, but don't take too long.

Passenger: OK, I'll be right back. (Pause)

Man on Bus: Mmm...Hmmm... What happened?

Janice: Don't try to sit up yet. Keep your head down. Are you feeling better?

Man: Yeah, but I feel a little shaky. I guess I passed out.

Janice: Yeah, you did, but don't try to talk. Just be quiet, 'til you feel better.

Passenger: Oh, I see he's coming around. Good. The driver says we'll be stopping in a minute. He'll feel better then.

OBJ. C-1

			Percentage
OBJ. B-2	19.	What happened to the man on the bus?	
		A. He stood up and bumped his head.	1
		B. He stopped breathing.	---
		C. He had a mild heart attack.	---
		*D. He fainted.	99
OBJ. B-3	20.	Why did the man on the bus feel queasy?	
		A. Because the air was polluted	2
		B. Because he hadn't eaten all day	2
		C. Because he was taking medication	1
		*D. Because the bus was shaking back and forth	95
OBJ. B-3	21.	What was the first thing the girl and the woman did to help the man?	
		A. They opened the window on the bus.	2
		*B. They lowered the man's head between his knees.	97
		C. They shook the man by his shoulders.	1
		D. They asked the man if he wanted a drink of water.	1†
OBJ. C-2	22.	What probably happened after the woman talked to the driver?	
		A. The bus slowed down until the man felt better.	13
		B. The man passed out again.	1
		C. The woman called the police.	1
		*D. The bus driver stopped as soon as possible.	86†

19

Results by Groups of Students

The results of students' performance on the listening test can also be reported in terms of various groups of students. Of particular interest are the results of students' performance by kind of community, sex, racial/ethnic and socioeconomic background characteristics. These categories are helpful to schools that wish to compare performance of their own students with the performance of students having similar characteristics. The results show that there are no statistically significant differences in the performance of males and females, but there are significant differences in the performance of students from various kind-of-community, racial/ethnic and socioeconomic groups. (See Table 3.). Statistically significant differences are those that cannot be accounted for by sampling variation. Even though statistically significant results have been identified, it should be noted that all of these groups of students performed well on the test.

Within the kind-of-community categories, students from residential suburbs performed significantly higher and students from big cities and industrial suburbs performed significantly lower than students statewide. The score for students from residential suburbs is about 2% above the statewide average. The scores for students from big cities and from industrial suburbs are about 3% and 2% below the statewide average, respectively.

Within racial/ethnic groups, white students performed significantly higher than students statewide, and all other racial/ethnic groups performed significantly lower. The score for white students is about 1% higher than the statewide average. The scores for black and Hispanic students are about 7% below the statewide average. The score for other minority students is about 4% below the statewide average.

Within socioeconomic (SES) groups students from higher SES backgrounds performed significantly higher than students statewide and students from lower SES backgrounds performed significantly lower than students statewide. The score for higher SES students is about 2% higher than the statewide average. The score for lower SES students is 3% lower than students statewide.

Results reported by groups of students must be interpreted with caution. Because there are fewer students in each group, these results are less precise than those previously reported. Furthermore, in some cases, small differences are statistically significant, and in other cases, significant differences are based on very small numbers of students. All these factors should be kept in mind in judging the educational significance of results. Also, note that one cannot attribute cause/effect interpretations to the results. The results merely suggest the relationships between various characteristics and performance.

TABLE 3. Average Percentage of Items Answered Correctly on the Total Listening Test for All Students and by Various Groups of Students

	Average Percentage Correct
ALL STUDENTS	88
KIND OF COMMUNITY	
Big City	85*
Industrial Suburb	86*
Residential Suburb	90*
Other	89
SEX	
Male	88
Female	88
RACIAL/ETHNIC BACKGROUND	
Black	82*
Hispanic	81*
White	89*
Other	84*
SOCIOECONOMIC BACKGROUND	
Lower	85*
Middle	88
Higher	90*

*Asterisks indicate statistically significant differences.

Other Available Information

More detailed results of the listening assessment are provided in the Massachusetts Basic Skills Assessment 1979-80: Technical Report. Included in that report are:

- Mean percentages and standard errors for the total test by major reporting groups (KOC, sex, racial/ethnic, white and minority, and bilingual and not bilingual)
- Mean percentages and standard errors for each skill by major reporting groups
- Mean percentages and standard errors for each situation by major reporting groups
- Mean percentages and standard errors for the total test by all reporting groups defined by the student questionnaire

Information about the development, reliability and validity of the listening test is included in the Massachusetts Basic Skills Assessment 1979-80: Development Report.

III. SPEAKING

Objectives and Instrument Development

The speaking assessment was designed to measure the 14 speaking objectives identified by the Massachusetts Basic Skills Improvement Policy. (See Exhibit 5.) Similar to the listening objectives, some of the speaking objectives deal with general speaking skills that apply to all speaking situations -- for example, Objective A-1, Use Words and Phrases Appropriate to the Situation. Other objectives deal with specific speaking situations -- for example, Objective C-1, Use Survival Words to Cope With Emergency Situations.

A set of speaking rating scales were developed for the speaking assessment. These scales were used in two ways:

- For classroom observation by teachers
- For one-on-one assessment by trained raters

The following sections describe the speaking rating scales, the teacher observation approach and the one-on-one approach.

Speaking Rating Scales

A set of four scales was developed to measure the objectives that deal with general speaking skills. Each scale measures one dimension of speaking skills. The dimensions include:

- Delivery
- Organization
- Content
- Language

The delivery dimension is concerned with how well a student transmits messages. It measures how well a student uses appropriate volume, rate and articulation while speaking. The organization dimension focuses on how well a student organizes messages. It measures how well a student sequences or shows relationships among ideas in messages. The content dimension is concerned with how well a student provides an adequate amount of relevant information to meet the requirements of various speaking tasks. In addition, it measures how well a student adapts the content of messages to specific listeners and situations. The language rat-

ing focuses on how well a student uses appropriate grammar and vocabulary while speaking.

EXHIBIT 5. Speaking Objectives

A. BASIC ORAL COMMUNICATION SKILLS

1. Use words and phrases appropriate to the situation
2. Speak loudly enough to be heard by a listener or group of listeners
3. Speak at a rate the listener can understand
4. Say words distinctly

B. PLANNING, DEVELOPING, AND STATING SPOKEN MESSAGES

1. Use words in an order that clearly expresses the thought
2. Organize main ideas for presentation
3. State main ideas clearly
4. Support main ideas with important details
5. Demonstrate knowledge of standard English usage

C. COMMON USES OF SPOKEN MESSAGES

1. Use survival words to cope with emergency situations
2. Speak so listener understands purpose
3. Ask for and give straightforward information
4. Describe objects, events and experiences
5. Question others' viewpoints

Within each speaking dimension, performance is rated using a 4-point scale:

1 = Inadequate
2 = Minimal
3 = Adequate
4 = Superior

The dimensions and the levels of the rating scales are explained in "The Speaking Assessment Ratings Guide." (See Appendix.)

The Teacher Observation Approach

The teacher observation approach is a general measure of a student's speaking performance. In this approach, two teachers who had the same student currently enrolled in class independently rated the student's general speaking performance in class using the speaking rating scales. Usually one of the teachers was an English teacher and the other from another subject area where students participate in a fair amount of classroom discussion and interaction, such as history, government or science.

Teachers based their ratings of a student on their observation of the student's performance in normal classroom activities, such as asking questions, responding to questions, explaining how to do something, giving a report to the class or talking with other students in discussion groups. They considered the student's average performance since the beginning of the semester.

The One-on-One Approach

The one-on-one approach is a focused measure of a student's speaking performance. In this approach, a trained rater rated the student's performance using the same speaking rating scales used by teachers for the teacher observation approach. However, instead of basing ratings on classroom observations, a rater gave a student specific tasks and rated the student's performance on these tasks.

The rater assessed each student individually. The rater gave the student four speaking tasks and rated the student's performance on each task along the four dimensions. Thus, for each task, the rater gave the student a rating from 1 to 4 for delivery, organization, content and language. The rating was assigned immediately after the student's response.

The speaking tasks used in the one-on-one approach reflect the objectives that deal with specific speaking situations. The tasks include:

- A description
- An emergency
- A sequence
- A persuasion task

Each task was familiar to all students and did not require any special knowledge or experience. The review committees judged that the tasks did not reflect sex, racial/ethnic or economic bias.

Future District Use of the Speaking Measures

The rationale for developing two approaches for assessing speaking is based on the need for measures that are feasible for large-scale use by school districts in the future. The intention is to use the teacher observation approach as a screening measure and to use the one-on-one approach as a back-up measure in cases where a student's level of ability is in question. School districts would use the teacher observation measure to assess all students in a grade. They would use the one-on-one measure to assess students whose level of ability is in question. For example, the one-on-one measure might be used when two teachers do not agree in their observation ratings of a student.

Before final plans are made for implementing the speaking measures in school districts, it is necessary to demonstrate that the two approaches are reliable and valid. Initial tests of the reliability and validity of the measures were conducted during the statewide assessment. The review committees felt that initial results were encouraging. However, further study will be undertaken before final recommendations are made regarding the measures. The results of the reliability and validity of the two approaches are reported in detail in the Massachusetts Basic Skills Assessment 1979-80: Development Report.

Results of Teacher Observation and One-on-One Ratings

Interpreting Results

The basic measure of speaking performance for both the teacher observation and one-on-one measures is a rating scale that ranges from 1 to 4, 1 designating the lowest rating and 4 designating the highest rating. This chapter contains the results for each teacher observation and for each one-on-one rating. For each rating, the report indicates the percentage of students who received ratings at each level of the rating scale, 1 through 4, and students' average, or mean, ratings. Ratings are also summarized in various ways. Average ratings are reported for each measure overall and by dimension and by task. Ratings for each measure are added together and reported as total scores. The teacher observation results only reflect those instances where a student

received ratings from two teachers and the ratings on each dimension are within one point of each other.

The speaking results presented in this chapter are based on the performance for a sample of 12th graders. They are estimates of the "true" performance that would have been obtained if all 12th graders in Massachusetts had participated in the speaking assessment. Statistical computations, standard errors, indicate that the estimates of average ratings are within .1 of the "true" ratings of the population. For example, if the average rating of students from the sample is 2.8, one can feel confident that the "true" rating of all students is between 2.7 and 2.9. Thus, the reader should not over emphasize the importance of small differences in ratings.

The dimensions of speaking skills used for both the teacher observation and one-on-one measures and the specific tasks used for the one-on-one measure are only a sample of all possible dimensions and tasks that could have been used. Thus, the reader should be cautious in making generalizations about the speaking results.

Results of the Teacher Observation Ratings

The average rating of students on the teacher observation measure overall is 2.95.[3] The average falls just under the 3 rating, which reflects adequate performance.

Exhibit 6 shows the average ratings of students for the teacher observation measure overall and for each dimension. The averages for dimensions range from 2.91 in delivery to 2.99 in content. The range is quite restricted, indicating that the level of performance of students as assessed by the teacher observation ratings does not vary much from dimension to dimension. All of the average ratings tend to fall just below 3, the rating that reflects adequate performance.

---

[3]The standard error for the teacher observation measure is .03 points.

EXHIBIT 6. Average Rating for the Teacher Observation Measure Overall and by Each Speaking Dimension

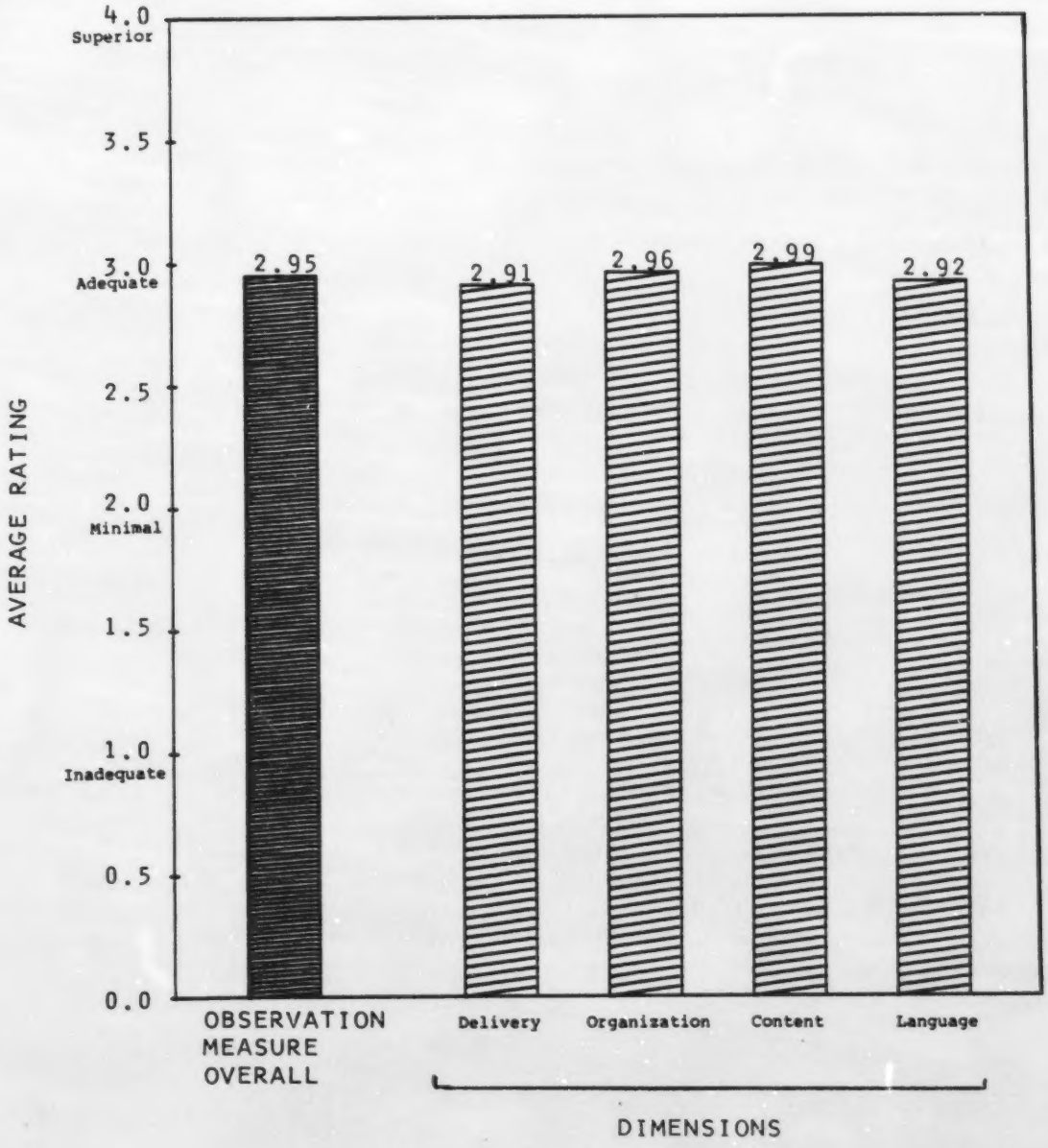


Table 4 indicates the percentage of students who received teacher observation ratings at each level of the rating scale, 1 through 4, for each dimension . The distribution of students across the four levels of the rating scale is about the same for delivery, organization, content and language. On the average, only about 1% of the students are rated inadequate, about 19% are rated minimal and 80% are rated adequate or superior.

TABLE 4. Percentage of Students Receiving Teacher Observation Ratings at Each Level for Each Dimension

	Percentage of Students			
	Inadequate 1	Minimal 2	Adequate 3	Superior 4
DELIVERY	1	20	65	13†
ORGANIZATION	1	18	67	15†
CONTENT	1	17	65	17
LANGUAGE	1	20	65	14
AVERAGE§	1	19	65	15

†Because of rounding errors, percentages might not total 100%.
§Average is the statistical average percentage of students receiving ratings at each level across the four dimensions. Students did not receive an average rating.

Another way to report the performance of students on the teacher observation measure is to add together their ratings from each teacher on each dimension to create a total teacher observation score. This score ranges from 8 (both teachers rate the student 1 on all four dimensions) to 32 (both teachers rate the student 4 on all four dimensions). Table 5 shows the distribution of students' total teacher observation scores. A score of 8 indicates that the students received all inadequate ratings; a score of 16 indicates minimal ratings on the average; a score of 24 indicates adequate ratings on the average; and a score of 32 indi-

cates all superior ratings. The largest percentage of students received scores in the range of 22 to 25. Only about 4% of the students received scores of 16 or lower.

TABLE 5. Distribution of Total Teacher Observation Scores

	Total Score	Percentage of Students	Cumulative Percentage
INADEQUATE	8	0	0
	9	0	0
	10	0	0
	11	--	--
	12	--	--
	13	0	--
	14	--	1
	15	1	2
MINIMAL	16	2	4
	17	2	5
	18	4	9
	19	4	13
	20	6	19
	21	6	25
	22	9	34
	23	13	47
ADEQUATE	24	18	65
	25	8	73
	26	7	79
	27	6	86
	28	5	91
	29	4	95
	30	3	98
	31	1	99
SUPERIOR	32	1	100

--Dashes are used for percentages which are less than .5.

Results of the One-on-One Ratings

The average rating of students on the one-on-one measure overall is 2.89.[4] The average rating for the one-on-one measure is lower than the average rating for the teacher observation measure (2.89 as compared with 2.95), but it still falls just under the 3 rating, which reflects adequate performance.

Exhibit 7 shows the average ratings of students for the one-on-one measure overall and for each dimension. In this case, the average for a dimension is a summary of the ratings of students within a dimension across four tasks. For example, the average rating for delivery is a summary of students' ratings on delivery for the four tasks. The averages for dimensions range from 2.85 in content to 2.94 in language. As with the teacher observation ratings, the range of the one-on-one ratings is quite restricted, indicating that the level of performance of students as assessed by the one-on-one measure does not vary much from dimension to dimension. All of the average ratings fall just below 3, which reflects adequate performance.

Exhibit 8 shows the average ratings of students for the one-on-one measure overall and for each task. In this case, the average for a task is a summary of the ratings of students within a task across the four dimensions. For example, the average rating for the description task is a summary of students' ratings on the four dimensions of the description task. The averages for tasks range from 2.84 on the description task to 2.97 on the emergency task. Again, the range is quite restricted, indicating that the level of performance of students as assessed by the one-on-one measure does not vary much from task to task. All of the average ratings fall just below 3, the rating that reflects adequate performance.

Tables 6 through 9 show the percentage of students who received one-on-one ratings at each level of the rating scale, 1 through 4, for each task. The actual task is indicated at the top of each table. The distribution of students across the four levels of the rating scale is about the same for all four tasks. Students are somewhat more likely to receive low ratings (1 and 2) in the organization and content dimensions than the delivery and language dimensions. On the average, very few students are rated inadequate, only about 1%. About 14% of the students are rated minimal, and about 85% of the students are rated adequate or superior, on the average.

---

[4]The standard error for the teacher observation measure is .02 points.

EXHIBIT 7. Average Rating for the One-on-One Measure Overall and by Each Speaking Dimension

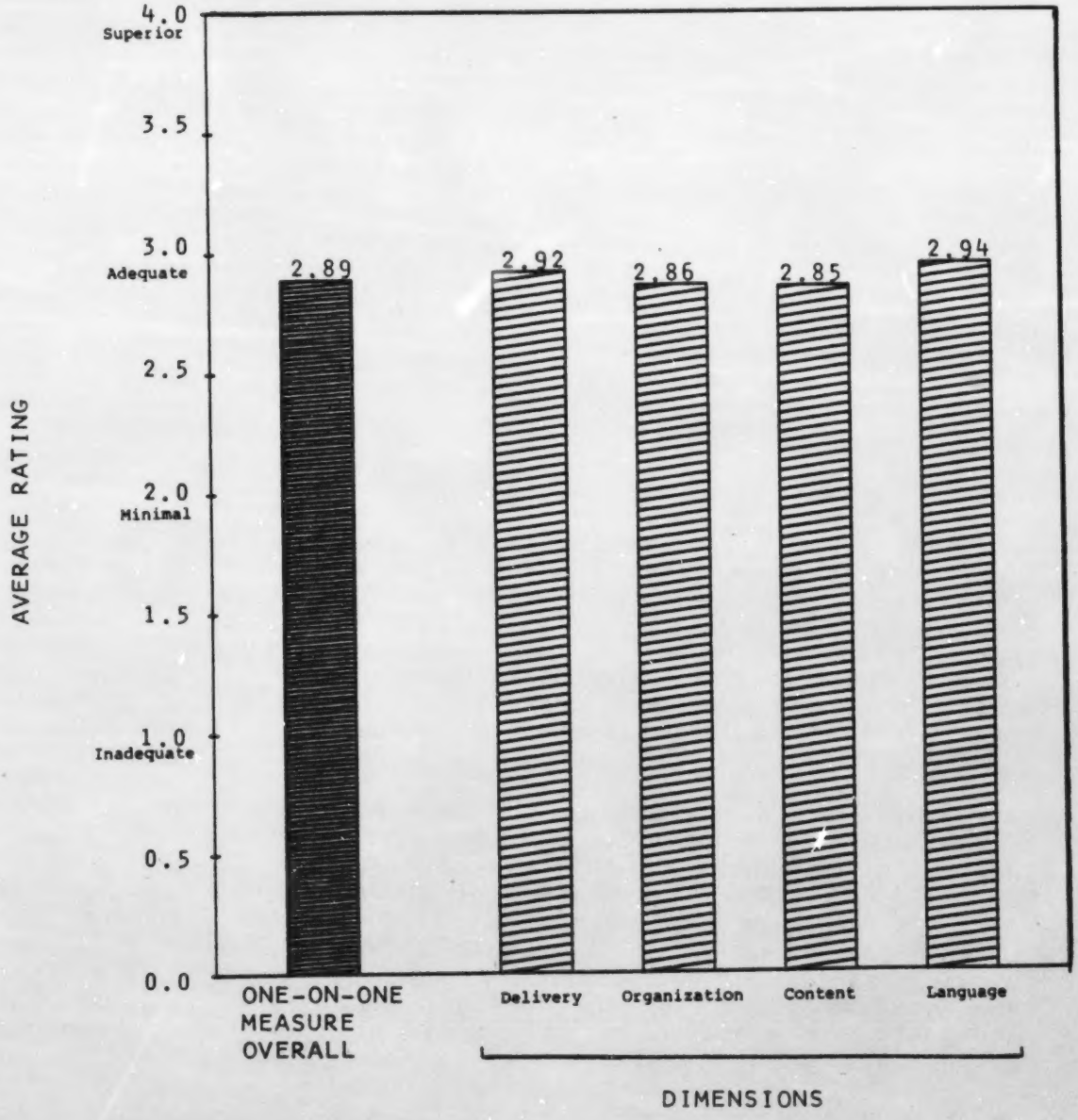


EXHIBIT 8. Average Rating for the One-on-One Measure Overall and by Each Speaking Task

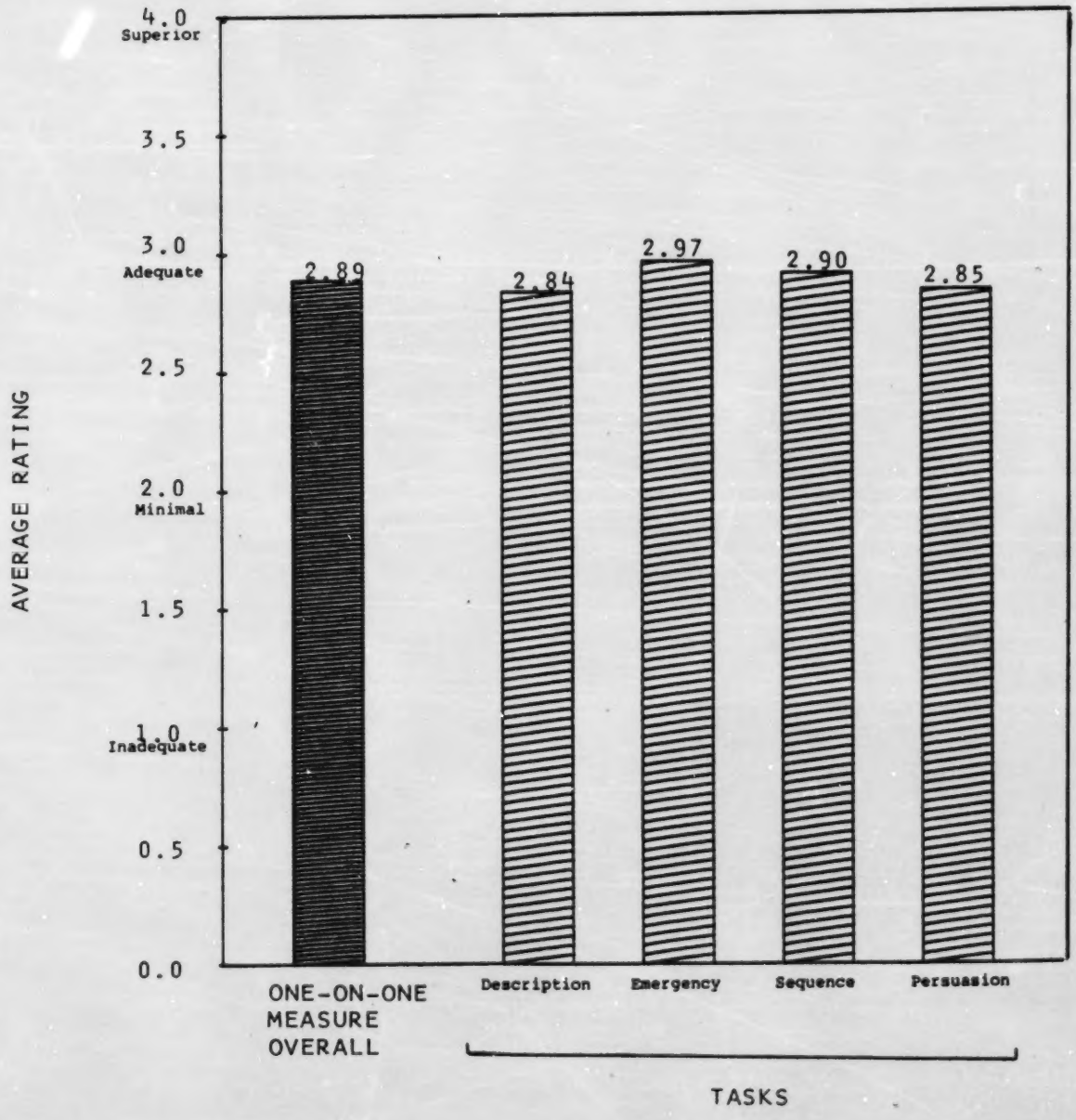


TABLE 6. Percentage of Students Receiving One-on-One Speaking Ratings at Each Level for Each Dimension of the Description Task

DESCRIPTION TASK:

THINK ABOUT YOUR FAVORITE CLASS OR EXTRACURRICULAR ACTIVITY IN YOUR SCHOOL. DESCRIBE TO ME EVERYTHING YOU CAN ABOUT IT SO THAT I WILL KNOW A LOT ABOUT IT. (HOW ABOUT SOMETHING LIKE A SCHOOL SUBJECT, A CLUB OR A SPORTS PROGRAM.)

	Percentage of Students			
	Inadequate 1	Minimal 2	Adequate 3	Superior 4
DELIVERY	1	10	87	2
ORGANIZATION	2	22	72	4
CONTENT	2	27	61	10
LANGUAGE	1	10	84	4†
AVERAGE§	2	17	76	5

†Because of rounding errors, percentages might not total 100%.
§Average is the statistical average percentage of students receiving ratings at each level across the four dimensions. Students did not receive an average rating.

TABLE 7. Percentage of Students Receiving One-on-One Speaking Ratings at Each Level for Each Dimension of the Emergency Task

EMERGENCY TASK:

IMAGINE THAT YOU ARE HOME ALONE AND YOU SMELL SMOKE. YOU CALL THE FIRE DEPARTMENT AND I ANSWER YOUR CALL. TALK TO ME AS IF YOU WERE TALKING ON THE TELEPHONE. TELL ME EVERYTHING I WOULD NEED TO KNOW TO GET HELP TO YOU. (TALK DIRECTLY TO ME. BEGIN BY SAYING HELLO.)

	Percentage of Students			
	Inadequate 1	Minimal 2	Adequate 3	Superior 4
DELIVERY	1	9	85	5
ORGANIZATION	1	13	74	12
CONTENT	1	17	68	14
LANGUAGE	1	6	85	9†
AVERAGE§	1	11	78	10

†Because of rounding errors, percentages might not total 100%.
§Average is the statistical average percentage of students receiving ratings at each level across the four dimensions. Students did not receive an average rating.

TABLE 8. Percentage of Students Receiving One-on-One Speaking Ratings at Each Level for Each Dimension of the Sequence Task

SEQUENCE TASK:

THINK ABOUT SOMETHING YOU KNOW HOW TO COOK. EXPLAIN TO ME STEP BY STEP HOW TO MAKE IT. (HOW ABOUT SOMETHING LIKE POPCORN, A SANDWICH OR EGGS.)

	Percentage of Students			
	Inadequate 1	Minimal 2	Adequate 3	Superior 4
DELIVERY	--	8	90	1†
ORGANIZATION	1	13	81	4†
CONTENT	1	17	75	7
LANGUAGE	1	7	90	1†
AVERAGE§	1	12	84	4†

--Dashes are used for percentages that are less than .5.
†Because of rounding errors, percentages might not total 100%.
§Average is the statistical average percentage of students receiving ratings at each level across the four dimensions. Students did not receive an average rating.

TABLE 9. Percentage of Students Receiving One-on-One Speaking Ratings at Each Level for Each Dimension of the Persuasion Task

PERSUASION TASK:

THINK ABOUT ONE CHANGE YOU WOULD LIKE TO SEE MADE IN YOUR SCHOOL, LIKE A CHANGE IN RULES OR PROCEDURES. IMAGINE I AM THE PRINCIPAL OF YOUR SCHOOL. TRY TO CONVINCE ME THAT THE SCHOOL SHOULD MAKE THIS CHANGE. (HOW ABOUT SOMETHING LIKE A CHANGE IN THE RULES ABOUT HALL PASSES OR THE PROCEDURES FOR ENROLLING IN COURSES.)

	Percentage of Students			
	Inadequate 1	Minimal 2	Adequate 3	Superior 4
DELIVERY	2	8	86	4
ORGANIZATION	2	19	76	3
CONTENT	2	24	67	7
LANGUAGE	2	9	85	4
AVERAGE§	2	15	78	4†

†Because of rounding errors, percentages might not total 100%.
§Average is the statistical average percentage of students receiving ratings at each level across the four dimensions. Students did not receive an average rating.

Another way to report performance of students on the one-on-one measure is to add together their ratings on each dimension for each task to create a total one-on-one score. This score ranges from 16 (a student receives ratings of 1 on all four dimensions for all four tasks) to 64 (a student receives ratings of 4 on all four dimensions for all four tasks). A score of 16 indicates the student received all inadequate ratings; a score of 32 indicates minimal ratings on the average; a score of 48 indicates adequate ratings on the average; and a score of 64 indicates all superior ratings. Table 10 shows the distribution of students' total one-on-one scores. The largest percentage of students received scores in the range of 44 to 50. Only about 2% of the students received scores of 32 or lower.

Speaking Results by Groups of Students

The results of student performance on the teacher observation and one-on-one speaking measures can also be reported in terms of various groups of students. Of particular interest are the results of students' performance by kind-of-community, sex, racial/ethnic and socioeconomic background characteristics. These characteristics are helpful to schools that wish to compare performance of their own students with the performance of students having similar characteristics. The results of the teacher observation measure indicate that there are no statistically significant differences in the performance of students except for one socioeconomic group. The results of the one-on-one measure indicate no statistically significant differences in the performance of groups of students except for one kind-of-community group and one socioeconomic group. (See Table 11.) Statistically significant differences are those that cannot be accounted for by sampling variation. Even though a few statistically significant differences have been identified, it is important to note that all of these groups of students performed well on both measures.

For the teacher observation measure, students from higher socioeconomic backgrounds performed significantly higher than students statewide. The average ratings for these students is .07 points above the statewide average.

For the one-on-one measure, students from big cities performed significantly lower than students statewide. The average rating for these students is .09 points below the statewide average. Also, students from middle socioeconomic backgrounds performed significantly lower than students statewide. The average rating for these students is .06 points below the state average.

TABLE 10. Distribution of Total One-on-One Scores

	Total Score∫	Percentage of Students	Cumulative Percentage
INADEQUATE	16	0	0
	18	0	0
	20	--	--
	22	0	--
	24	--	--
	26	--	1
	28	--	1
	30	--	1
MINIMAL	32	1	2
	34	1	3
	36	1	4
	38	2	6
	40	4	10
	42	4	14
	44	12	25
	46	17	43
ADEQUATE	48	34	77
	50	12	89
	52	5	94
	54	3	97
	56	2	99
	58	--	39
	60	1	100
	62	0	100
SUPERIOR	64	0	100

∫Scores are collapsed into intervals of 2.
--Dashes are used for percentages which are less than .5.

TABLE 11. Average Teacher Observation Ratings and Average One-on-One Ratings for All Students and by Various Reporting Groups

	Average Teacher Observation Rating	Average One-on-One Rating
ALL STUDENTS	2.95	2.89
KOC		
Big City	3.01	2.80*
Industrial Suburb	2.88	2.87
Residential Suburb	2.98	2.91
Other	2.87	2.94
SEX		
Male	2.89	2.87
Female	2.99	2.91
RACIAL/ETHNIC		
Black	2.94	2.92
Hispanic	2.76	2.75
White	2.95	2.89
Other	3.04	2.92
SES		
Lower	2.82	2.84
Middle	2.87	2.83*
Higher	3.02*	2.95

*Asterisks indicate statistically significant differences.

Results reported by groups of students must be interpreted with caution. Because there are fewer students in each group, these results are less precise than those previously reported. Furthermore, in some cases, small differences are statistically significant and in other cases, significant differences are based on very small numbers of students. All these factors should be kept in mind in judging the educational significance of results. Also, note that one cannot attribute cause/effect interpretations to the results. The results merely suggest the relationships between various characteristics and performance.

Other Available Information

More detailed results of the speaking assessment are provided in the Massachusetts Basic Skills Assessment 1979-80: Technical Report. Included in that report are:

- Mean ratings and standard errors for the teacher observation measure overall by major reporting groups (KOC, sex, racial/ethnic, white and minority, and bilingual and not bilingual)
- Mean ratings and standard errors for each dimension of the teacher observation measure by major reporting groups
- Mean ratings and standard errors for the teacher observation measure overall by all reporting groups defined by the student questionnaire
- Mean ratings and standard errors for the one-on-one measure overall by major reporting groups
- Mean ratings and standard errors across each dimension of the one-on-one measure by major reporting groups
- Mean ratings and standard errors across each task of the one-on-one measure by major reporting groups
- Mean ratings and standard errors for each dimension of each task of the one-on-one measure by major reporting groups
- Mean ratings and standard errors for the one-on-one measure overall by all reporting groups defined by the student questionnaire

Information about the development, reliability and validity of the teacher observation and one-on-one measures are included in the Massachusetts Basic Skills Assessment 1979-80: Development Report.

IV. BACKGROUND INFORMATION

Introduction

The main purpose of the statewide assessment is to collect information on the listening and speaking performance of 12th graders. A secondary purpose of the assessment is to collect background information that provides additional insight into the performance of students and baseline information prior to the implementation of the Basic Skills Improvement Policy.

Background information was collected through questionnaires administered to principals, teachers and students participating in the assessment. Some of the questions focused on the nature of the school environment, including general descriptive information and specific aspects of school programs that might influence listening and speaking skills. Other questions focused on characteristics of the student, including general descriptive information and specific factors that might impact listening or speaking skills.

Some of the background questions asked for factual information, for example, "What percent of the 12th graders in your school are not expected to receive a high school diploma in June?" Some of the background questions asked for opinions, for example, "What is your opinion of the amount of emphasis on listening and speaking skills instruction in your school?" All of the background questions rely upon the self-reported responses of principals, teachers and students. Thus, the data reflect the perceptions and opinions of the individuals responding to the questions.

School Characteristics

Most of the information regarding school characteristics was obtained from the Principal Questionnaire. Forty-four out of 49 participating principals responded to this questionnaire. Unlike the student results, responses of principals reflect only the characteristics of the schools participating in the assessment and cannot be generalized to all secondary schools in Massachusetts. This is because the sampling procedures used in the assessment were designed to be representative of students but not of schools.

Curriculum

An important background factor related to listening and speaking performance is the amount of instruction and extracurricular activities focusing on these skills. Principals were asked about the presence of various topics related to speaking and listening in their curriculum. These topics include: (1) debate, forensics or argumentation; (2) discussion or group process; (3) interpersonal communication or human relations; (4) speech or public speaking; (5) radio, TV or film; and (6) theater, drama or oral interpretation. Fifty-seven percent of the principals reported that they include all six topics in the curriculum either as a concentrated unit of study of nine weeks or more or as a topic integrated into a general course. One principal reported that none of these topics is included in the curriculum. More than half the principals reported having a concentrated unit of study in speech. Similarly, more than half reported having a concentrated unit in theater. About a third of the principals reported that debate is not included in the school's curriculum. Few schools have discussion or interpersonal communication included in their curriculum as a concentrated unit of study, but many schools have these topics integrated into a general course.

Extracurricular Activities

The principals were also asked about the presence of extracurricular activities related to listening and speaking skills. These activities include: (1) debate or forensics; (2) public speaking; (3) radio, TV or film; and (4) theater or drama. Sixteen percent of the principals reported that they have all four activities included in their extracurricular programs. Seven percent of the principals reported that they do not include any of these activities in their extracurricular programs. The most popular extracurricular activity among the four is theater; 93% of the principals reported having this activity. All the other activities were reported as present by approximately one-third to one-half of the principals.

Emphasis on Listening and Speaking

Principals were asked their opinion of the amount of emphasis on listening and speaking skills instruction in their schools. Sixty-four percent of the principals believe that there is insufficient instruction in the area of listening, and 69% of the principals believe that there is insufficient instruction in the area of speaking.

Students overwhelmingly reported that they have received limited instruction in listening and speaking skills. Over 90% of the students participating in the listening assessment reported that they were rarely given work in school aimed at improving their listening skills. About 90% of the students participating in the speaking assessment reported that they were rarely given work in school aimed at improving their speaking skills. Also, more than half (64%) of the students in the speaking assessment have never participated in extracurricular or community activities that emphasize speaking skills.

Both principals and students expressed their opinion with respect to the level of speaking and listening skills of students leaving high school. Fifty-three percent of the principals believe that students can graduate without knowing how to listen and speak very well. Students felt about the same way -- 44% of the students in the listening assessment and 49% of the students in the speaking assessment.

These data indicate that most schools include some listening and speaking skills in the curriculum and extracurricular program. However, principals and students alike perceive a deficit in the skills of many students and a need for more instruction.

Teachers

Another area of special concern is the reaction of teachers to the teacher observation rating measure. This information was collected from 591 teachers participating in the teacher observation ratings who completed the Teacher Questionnaire. About 45% of the participating teachers taught English; about 25% taught social studies; about 30% taught other subjects. The responses of English and speech teachers were very similar to the responses of other teachers, indicating that all teachers reacted similarly to the teacher observation measure. Most of the teachers (86%) had no difficulty understanding the "Speaking Assessment Ratings Guide." About 90% of the teachers were confident about the accuracy of the observation ratings they made. The review committees recommended that all teachers involved in the teacher observation measure receive specific training in the use of the speaking ratings scales.

Student Characteristics

A limited amount of data was collected on student characteristics. Some of this information relates to factors that might influence listening and speaking skills. Information was collected concerning the students' first language. Only 6% of the students

in the listening assessment and 5% of the students in the speaking assessment were not native English speakers. Information was also collected concerning hearing and speech problems of students. Only 4% of the students in the listening assessment reported that they had ever received help for a hearing problem, and 15% of the students in the speaking assessment reported that they had ever received help for a speech difficulty. Only 3% of the students in the listening assessment reported having difficulty understanding things they needed or wanted to listen to fairly or very often. About 11% of the students in the speaking assessment reported having difficulty talking about things they needed or wanted to talk about fairly or very often. The results indicate that very few students in the listening and speaking assessment displayed characteristics that might negatively influence performance. However, in almost all cases these students performed statistically significantly lower than students statewide.

Most students in the assessment believe that all students should be able to demonstrate basic listening and speaking skills by the time they finish high school. Eighty-three percent of the students in the listening assessment feel that students should be required to learn the skills measured by the listening assessment. Ninety-two percent of the students in the speaking assessment feel that students should be able to do the tasks they were asked to do in the one-on-one measure.

Other Available Information

Full results of the Principal Questionnaire, Teacher Questionnaire and Student Questionnaires are presented in the Massachusetts Basic Skills Assessment 1979-80: Technical Report. Included in the report are:

- Percentage of principals responding to each option of each question
- Percentage of teachers responding to each option of each question
- Percentage of students responding to each option of each question

APPENDIX

MASSACHUSETTS DEPARTMENT OF EDUCATION ASSESSMENT OF BASIC SKILLS

SPEAKING ASSESSMENT RATINGS GUIDE

OVERVIEW

There are numerous kinds of speaking tasks that students must perform in everyday life, both in school and out of school. The Massachusetts Basic Skills Improvement Policy has focused on some of these tasks, including describing objects, events and experiences, explaining the steps in a sequence, providing information in an emergency and persuading someone.

In order to accomplish a speaking task, the speaker must formulate and transmit a message to a listener. This process involves deciding what needs to be said, organizing the message, adapting the message to the listener and situation, choosing language to convey the message and finally delivering the message. The effectiveness of the speaker may be rated in terms of how well the speaker meets the requirements of the task.

The Massachusetts test of basic skills in speaking separates speaking skills into four dimensions:

Delivery
Organization
Content
Language

Delivery is concerned with the transmission of the message, i.e., volume, rate and articulation. Organization is concerned with how the content of the message is sequenced and how the ideas are related to one another. Content is concerned with the amount and relevance of information in the message, and how the content is adapted to the listener and situation. Language is concerned with the grammar and words which are used to convey the message.

Each of the four dimensions is rated on a four point scale: 1 is the lowest rating and 4 is the highest rating. A general set of principles underlies the rating scale for all four components. Ratings of 1 reflect speaking skills which are inadequate in meeting the requirements of the task. Ratings of 2 reflect speaking skills which are minimal in meeting the requirements of the task. Ratings of 3 reflect speaking skills which are adequate in meeting the requirements of the task. Ratings of 4 are superior in meeting the requirements of the task.

Individuals who act as raters for the speaking assessment need to take the role of a naive, objective listener. The rater must be naive so that the rater can base his or her rating on exactly what the speaker says. The rater must be careful not to let his or her own knowledge and experience influence the rating. The rater must face each speaker as if it were a new experience. The rater must also be objective so that he or she does not let a particular set of norms of social acceptability influence the rating. The rater must evaluate the speaker in terms of how well the speaker meets the requirements of the speaking task, irrespective of the particular communication style the speaker uses.

DELIVERY

The delivery rating focuses on the transmission of the message. It is concerned with volume, rate and articulation. Articulation refers to pronunciation and enunciation. Some examples of poor articulation include mumbling, slurring words, stammering, stuttering and exhibiting disfluencies such as ahs, uhms or "you knows."

1 = The delivery is inadequate in meeting the requirements of the task.

e.g., The volume is so low that you cannot understand most of the message.
The rate is so fast that you cannot understand most of the message.
The pronunciation and enunciation are so unclear that you cannot understand most of the message.

2 = The delivery is minimal in meeting the requirements of the task.

e.g., The volume is too low or too loud.
The rate is too fast or too slow. Pauses are too long or at inappropriate spots.
The pronunciation and enunciation are unclear. The speaker exhibits many disfluencies such as ahs, uhms or "you knows."
You are distracted by problems in the delivery of the message.
You have difficulty understanding the words in the message. You have to work to understand the words.

3 = The delivery is adequate in meeting the requirements of the task.

e.g., The volume is not too low or too loud.
The rate is not too fast or too slow. Pauses are not too long or at inappropriate spots.
The pronunciation and enunciation are clear. The speaker exhibits few disfluencies, such as ahs, uhms and "you knows."
You are not distracted by problems in the delivery of the message.
You do not have difficulty understanding the words in the message.

4 = The delivery is superior in meeting the requirements of the task.

e.g., The speaker uses delivery to emphasize and enhance the meaning of the message. The speaker delivers the message in a lively, enthusiastic fashion.
The volume varies to add emphasis and interest.
Rate varies and pauses are used to add emphasis and interest.
Pronunciation and enunciation are very clear. The speaker exhibits very few disfluencies such as ahs, uhms or "you knows."

NOTE: In articulation you may be concerned with accent. However, articulation should be rated with respect to your ability to understand the message, not the social acceptability of the accent. One particular accent is not considered better than another. REMEMBER, in this component you are rating how the student speaks, not what the student says.

ORGANIZATION

The organization rating focuses on how the content of the message is structured. It is concerned with sequence and the relationships among the ideas in the message.

1 = The organization is inadequate in meeting the requirements of the task.

e.g., The message is so disorganized that you cannot understand most of the message.

2 = The organization is minimal in meeting the requirements of the task.

e.g., The organization of the message is mixed up; it jumps back and forth.
The organization of the message appears random or rambling.
You have difficulty understanding the sequence and relationships among the ideas in the message. You have to make some assumptions about the sequence and relationships of ideas.
You cannot put the ideas in the message into an outline.

3 = The organization is adequate in meeting the requirements of the task.

e.g., The message is organized.
You do not have difficulty understanding the sequence and relationships among the ideas in the message. You do not have to make assumptions about the sequence and relationships of ideas.
You can put the ideas in the message into an outline.

4 = The organization is superior in meeting the requirements of the task.

e.g., The message is overtly organized.
The speaker helps you understand the sequence and relationships of ideas by using organizational aids such as announcing the topic, previewing the organization, using transitions and summarizing.

NOTE: Make sure you are not unconsciously "filling in" organization for a speaker, because you happen to know something about the speaker's topic. If you have to make assumptions about the organization, this fact should be reflected in your rating. REMEMBER, in this component you are rating how the student organizes the message, not what the student says.

CONTENT

The content rating focuses on the specific things which are said. It is concerned with the amount of content related to the task, the relevance of the content to the task and the adaptation of the content to the listener and the situation.

1 = The content is inadequate in meeting the requirements of the task.

e.g., The speaker says practically nothing.
The speaker focuses primarily on irrelevant content.
The speaker is highly egocentric. The speaker appears to ignore the listener and the situation.

2 = The content is minimal in meeting the requirements of the task.

e.g., The speaker does not provide enough content to meet the requirements of the task.
The speaker includes some irrelevant content. The speaker wanders off the topic.
The speaker adapts poorly to the listener and the situation. The speaker uses words and concepts which are inappropriate for the knowledge and experiences of the listener (e.g., slang, jargon, technical language). The speaker uses arguments which are self-centered rather than other-centered.

3 = The content is adequate in meeting the requirements of the task.

e.g., The speaker provides enough content to meet the requirements of the task.
The speaker focuses primarily on relevant content. The speaker sticks to the topic.
The speaker adapts the content in a general way to the listener and the situation. The speaker uses words and concepts which are appropriate for the knowledge and experience of a general audience. The speaker uses arguments which are adapted to a general audience.

4 = The content is superior in meeting the requirements of the task.

e.g., The speaker provides a variety of types of content appropriate for the task, such as generalizations, details, examples and various forms of evidence.
The speaker adapts the content in a specific way to the listener and situation. The speaker takes into account the specific knowledge and experience of the listener, adds explanations as necessary and refers to the listener's experience. The speaker uses arguments which are adapted to the values and motivations of the specific listener.

NOTE: This rating is concerned with content in terms of quantity, relevance and adaptation. It is not concerned with content in terms of accuracy. Concerns with accuracy of content fall outside a speaking skills assessment. Also, make sure you are not unconsciously "filling in" content for a speaker because you happen to know something about the speaker's topic. If you add information, this fact should be reflected in your rating. REMEMBER, in this component you are rating the quantity, relevance and adaptation of what the student says, not the accuracy of what the student says.

LANGUAGE

The language rating deals with the language which is used to convey the message. It is concerned with grammar and choice of words.

1 = The language is inadequate in meeting the requirements of the task.

e.g., The grammar and vocabulary are so poor that you cannot understand most of the message.

2 = The language is minimal in meeting the requirements of the task.

e.g., The speaker makes many grammatical mistakes.
The speaker uses very simplistic, bland language. The speaker uses a "restricted code," a style of communication characterized by simple grammatical structure and concrete vocabulary.

3 = The language is adequate in meeting the requirements of the task.

e.g., The speaker makes few grammatical mistakes.
The speaker uses language which is appropriate for the task, e.g., descriptive language when describing, clear and concise language when giving information and explaining, persuasive language when persuading. The speaker uses an "elaborated code," a style of communication characterized by complex grammatical structure and abstract vocabulary.

4 = The language is superior in meeting the requirements of the task.

e.g., The speaker makes very few grammatical mistakes.
The speaker uses language in highly effective ways to emphasize or enhance the meaning of the message. As appropriate to the task, the speaker uses a variety of language techniques such as vivid language, emotional language, humor, imagery, metaphor, simile.

NOTE: In language you may be concerned with students who come from backgrounds where a foreign language or a non-standard form of English is spoken. However, language should be rated with respect to your ability to understand the message, not the social acceptability of the communication style. If a speaker's use of incorrect or non-standard English grammar interferes with your ability to understand the message, this fact should be reflected in your rating. REMEMBER, in this component you are rating how the student conveys the message through language, not what the student says.

ACKNOWLEDGMENTS

Advisory Committee on Basic Skills Improvement
1979-80

Bedros Kamitian (Chair)	Secondary School Principal	Holliston
Virginia Yardumian (V-Chair)	School Committee	Malden
William Fisher (Secretary)	Superintendent	Hamilton-Wenham
Fred Andelman	Mass. Teachers Association	Boston
Blair Arsenault	Secondary School Principal	Oakmont Regional
Virginia Barton	Parent	Lynn
Mildred Berkowitz	English Teacher	Dartmouth
Claudette Chagnon	Employer	Newton
Robert Coelho	Superintendent	Attleboro
Thomas Deveney	Administrator	Boston
Janet Doe	Secondary School Principal	Worcester
Roberta Doering	School Committee	Agawam
Donald Frizzle	Superintendent	Amherst-Pelham
Paul Gauvreau	Junior High Principal	Pittsfield
Walter Gleason	Assistant Superintendent	Winchester
Richard Granger	Junior High Principal	Auburn
Philip Hallorsn	Director of Mathematics	Springfield
Ann Keenan	Director of English	Braintree
John Kent	Elementary School Principal	Holbrook
Gail Latouf	Student	Boston
Ronald Laviolette	Elementary School Principal	Palmer
Steven Levine	Special Needs Coordinator	E. Bridgewater
Benjamin Levy	Doctoral Candidate	Somerville
Pamela Mason	Director of Reading	Reading
Gertrude Moskowitz	Parent	Pittsfield
Margarita Muniz	Bilingual Department	Boston
Edward Reidy	Administrator	Fitchburg
Virginia Ross	English Department Head	Marshfield
Charlotte Ryan	Mass.Parent-Teacher-Student Assoc.	Orange
Thomas Tanous	Guidance Counselor	Beverly
Nancy Thompson	Teacher	Boston
David Tobin	Superintendent	Assabet Valley Reg.
Yolanda Ulloa	Bilingual Education Director	Springfield
Harry Walen	Administrator (Retired)	Rockport
Philip Wormwood	Secondary School Principal	Andover

ACKNOWLEDGMENTS *(Continued)*

Communication and Listening/Speaking Skills Advisory Committees
1979-80

Anna Andruszkiewicz	Bilingual Teacher	Peabody
Maria Argento	High School Student	West Boylston
Kenneth Brown	Professor, Speech Communication	University of Massachusetts
Rhoda Caesar	Elementary School Teacher	Pittsfield
Joseph Dudrick	Director, Business Development	Sperry Research Center
Joanne Gurry	High School Teacher	Marshfield
Patricia Keating	High School Teacher	Northampton
Ann Keenan	Director of English	Braintree
Barbara Krysiak	Director of English	Waltham
Eileen Malloy	English Department Head	Whitman-Hanson
Pamela Mason	Director of Reading	Reading
William A. Merrill	Guidance Counselor	Lowell
Norman Najimy	Director of English	Pittsfield
Owen O'Neil	Director of Reading	Springfield
Bernice Randell	Parent Coordinator	Boston
Virginia Ross	English Department Head	Marshfield
Maureen Tisei	Communication Arts Dept. Head	Boston
John Tsang	Bilingual Department	Boston
Anne G. Welch	Middle School Teacher	Clinton
Jane Sherman Consultant to the Committee	Bureau of Equal Educational Opportunity	Springfield Office